# Hyperplane separation theorem and intro to convex analysis

Hedi Xia

June 27, 2017

## Theorem

Definition: If $C_1, C_2 \subseteq \mathbb{R}^n$ are nonempty, convex and disjoint ($C_1 \cap C_2 = \emptyset$), then there exists a hyperplane $H \subset \mathbb{R}^n$ such that $H$ separate $C_1$ and $C_2$.

- Convex
- Hyperplane
- Separate

## Convexity

A set $S$ is convex in $\mathbb{R}^n$ iff $\forall x, y \in S$, $l(x, y) \in S$. Where $l(x, y) = \{\lambda x + (1-\lambda)y : \lambda \in [0, 1]\}$.
Theorems about convexity: Given $C_1$ and $C_2$ convex,

- $C_1 \cap C_2$ is convex
- $C_1 + C_2$ is convex
- $C_1 + v$ is convex
- $AC_1$ is convex, where $A$ is a matrix

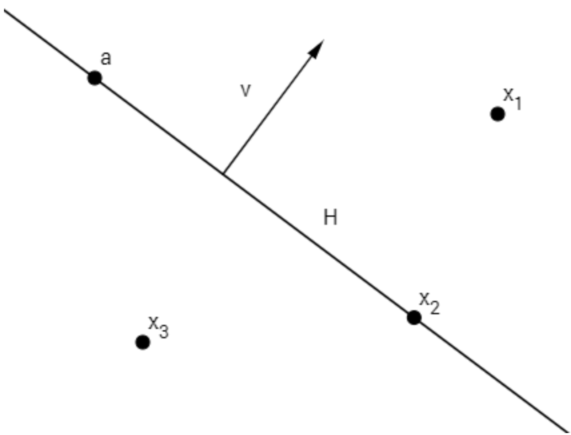
## Hyperplane

Definition: A hyperplane is a flat in $\mathbb{R}^n$ with dimension $n - 1$.
Definition for flat and its dimension:
$F$ is a flat of dimension $k$ iff $F = W + v$ for some $v \in \mathbb{R}^n$ and $W \leq \mathbb{R}^n$ such that $dim(W) = k$.
Theorems about Hyperplanes and flats:

- Every flat is convex, and so is hyperplanes and subspaces.

- A hyperplane can be represented by a perpendicular vector and a arbitrary point on the hyperplane. In this case, $H(v,a) = \{x \in \mathbb{R}^n : \langle v, x-a\rangle = 0\}$
- A hyperplane can also be represented by a linear functional and a scalar where: $H(f,t) = \{x \in \mathbb{R}^n : f(x) = t\}$

## Separation




Example in $\mathbb{R}^2$, we know that:

- $v \cdot (x_1 - a) > 0$
- $v \cdot (x_2 - a) = 0$
- $v \cdot (x_3 - a) < 0$

In general, hyperplane $H = (v,a)$ can separate the whole space into:

- $\{x \in \mathbb{R}^n : \langle v, x-a\rangle > 0\}$
- $\{x \in \mathbb{R}^n : \langle v, x-a\rangle = 0\} = H$
- $\{x \in \mathbb{R}^n : \langle v, x-a\rangle < 0\}$

We say that $H$ separates two half spaces defined as:

$$C_{\geqslant} = \{x \in \mathbb{R}^n : \langle v, x-a\rangle \geqslant 0\}$$

$$C_{\leq} = \{x \in \mathbb{R}^n : \langle v, x-a\rangle \leq 0\}$$

Notice that the second notation of hyperplane $H(f,t)$ is closely related with the first one with the following correspondence: $t = \lambda\langle v, a\rangle$ and $f(x) = \lambda\langle v, x\rangle$ where $\lambda$ is a nonzero real number. So we can also define:

$$C_{\geqslant} = \{x \in \mathbb{R}^n : f(x) \geqslant t\}$$

$$C_{\leq} = \{x \in \mathbb{R}^n : f(x) \leq t\}$$

We say that a hyperplane $H$ support a set $S$ if and only if either $S \subseteq C_{\geqslant}$ or $S \subseteq C_{\leq}$. Otherwise we say that $H$ cut $S$.

- If $H$ cut $S$, then $S \cap H \neq \emptyset$
- If a hyperplane $H$ supports a flat $F$, then $F \subseteq H$
- If $C$ is convex and $x \notin C$, then there exists a hyperplane $H$ such that $x \in H$ and $H$ support $C$. In addition, if $C$ is either closed or open, we can find one that is also disjoint with $C$. (Proof by induction on dimension)

## Proof

Because $C_1 \cap C_2 = \emptyset$, then $0_v \notin C_1 - C_2$. Notice that $C_1 - C_2$ is convex and $0_v \notin C_1 - C_2$, so there is a hyperplane $H$ such that $0_v \in H$ and $H$ support $C_1 - C_2$. So we can denote $H = (f, 0)$ in our second notion. So, either $f(C_1 - C_2) \leq 0$ or $f(C_1 - C_2) \geqslant 0$. Without loss of generality, suppose that $f(C_1 - C_2) \leq 0$. Then, $f(C_1) \leq f(C_2)$, so there is a lower bound for $f(C_2)$ because $f(C_1)$ is not empty. Let $m$ be the greatest lower bound, so $f(C_1) \leq m \leq f(C_2)$. This means that hyperplane $(f, m)$ separates $C_1$ and $C_2$.

## Usage

- Farkas' lemma: Let $A \in M_{m*n}(\mathbb{R})$ and $b \in \mathbb{R}^m$, then exactly one of the following two sets of equalities/inequalities has a solution,
  1. $Ax = b$ and $x \geqslant 0$
  2. $A^T y \geqslant 0$ and $b^T y < 0$
- Strong duality theorem in linear programming: If $P$ is a linear programming with an optimal value, then its dual has the same optimal value.

## Notation

- Let $A$,$B$ be set, then $A + B = \{a + b : a \in A, b \in B\}$
  Let $v$ be a vector, then $A + v = A + \{v\} = \{a + v : a \in A\}$
- Let $v$ be a vector, then we say that $v \geqslant 0$ if and only if every component of $v$ is nonnegative. We say that $v \leq 0$ if and only if every component of $v$ is nonpositive. We say that a vector $v_1 \leq v_2$ iff $v_1 - v_2 \leq 0$ and $v_1 \geqslant v_2$ iff $v_1 - v_2 \geqslant 0$